

PATENT ABSTRACTS OF JAPAN

(11)Publication number : **2001-296991**

(43)Date of publication of application : **26.10.2001**

(51)Int.Cl.

G06F 3/16

G10L 15/00

G10L 15/28

G10L 15/22

G10L 15/24

(21)Application number : **2000-112416**

(71)Applicant : **CANON INC**

(22)Date of filing : **13.04.2000**

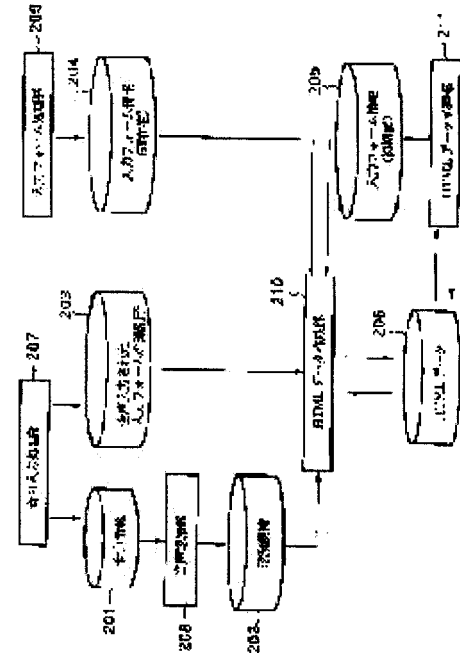
(72)Inventor : KUSHIDA AKIHIRO
KOSAKA TETSUO
ITO SHIRO

(54) DATA PROCESSING APPARATUS, DATA PROCESSING METHOD AND RECORDING MEDIUM

(57)Abstract:

PROBLEM TO BE SOLVED: To display and select a recognition candidate in sound input to an input form on a browser.

SOLUTION: This data processing apparatus is provided with a sound input processing part 207 for inputting information by sound to one of input sections included in the display contents of HTML data, a sound recognition part 208 for performing sound recognizing processing to the inputted sound to generate its recognition candidate and an HTML data generation part 210 detecting the input data given sound input among the HTML data and converting the HTML data with respect to the detected input section to HTML data for selecting and inputting the generated recognition candidate. The recognition candidate of sound recognition is displayed directly in the input section.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(51) Int.Cl. ⁷	識別記号	F I	ページ* (参考)
G 0 6 F 3/16	3 2 0	G 0 6 F 3/16	3 2 0 H 5 D 0 1 5
G 1 0 L 15/00		G 1 0 L 3/00	5 5 1 A
15/28			5 5 1 P
15/22			5 6 1 C
15/24			5 6 1 E
審査請求 未請求 請求項の数16 O L (全 12 頁) 最終頁に続く			

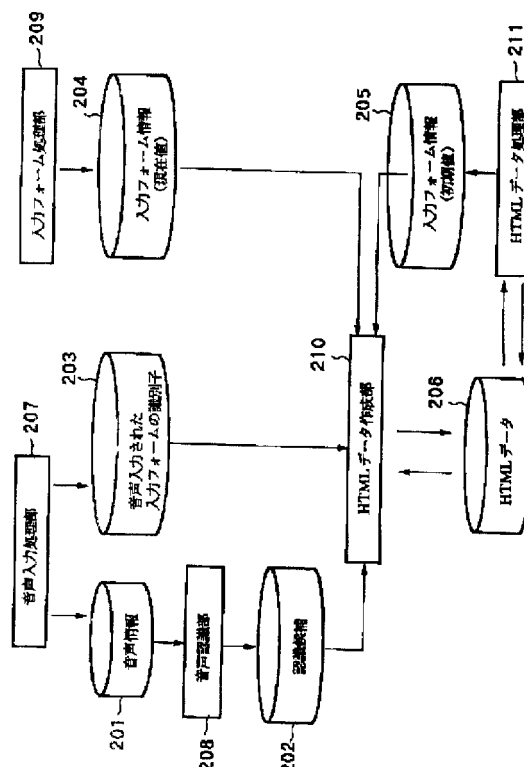
(21) 出願番号	特願2000-112416(P2000-112416)	(71) 出願人	000001007 キヤノン株式会社 東京都大田区下丸子3丁目30番2号
(22) 出願日	平成12年4月13日(2000. 4. 13)	(72) 発明者	梶田 晃弘 東京都大田区下丸子3丁目30番2号 キヤ ノン株式会社内
		(72) 発明者	小坂 哲夫 東京都大田区下丸子3丁目30番2号 キヤ ノン株式会社内
		(74) 代理人	100076428 弁理士 大塚 康徳 (外2名)
		最終頁に続く	

(54) 【発明の名称】 データ処理装置、データ処理方法、記録媒体

(57) 【要約】

【課題】 ブラウザ上の入力フォームに音声入力における認識候補を表示し、選択可能とすること。

【解決手段】 HTMLデータの表示内容に含まれる入力欄の1つに対して、音声により情報を入力するための音声入力処理部207と、入力された前記音声に対して音声認識処理を実行し、その認識候補を作成する音声認識部208と、HTMLデータのうち、音声入力になされた入力欄を検出し、検出した入力欄に対するHTMLデータを、作成した認識候補を選択して入力するHTMLデータに変換するHTMLデータ作成部210とを備え、入力欄に音声認識の認識候補が直接表示されるようにする。



【特許請求の範囲】

【請求項1】 所定のマークアップ言語で記述されたグラフィカルインターフェース用のデータの内容を表示する表示手段と、

前記表示手段による表示内容に含まれる入力欄の1つに対して、音声により情報を入力するための音声入力手段と、

入力された前記音声に対して音声認識処理を実行し、その認識候補を作成する音声認識手段と、

前記表示内容に含まれる入力欄のうち、前記音声入力手段により入力となされた入力欄を検出する検出手段と、前記検出手段により検出した入力欄に対する前記データを、作成した前記認識候補に基づいて変換する手段と、を備えたことを特徴とするデータ処理装置。

【請求項2】 前記検出手段により検出された入力欄が、該入力欄にテキストを入力するタイプの入力欄であった場合に、前記変換手段は、前記入力欄が、作成した前記認識候補を選択候補とする選択リストから候補を選択して入力するタイプとなるように、前記データを変換することを特徴とする請求項1に記載のデータ処理装置。

【請求項3】 前記音声認識手段は、複数の前記認識候補の各々の尤度を計算し、前記変換手段は、前記認識候補のリストが各候補の尤度の順番に表示されるように、前記データを変換することを特徴とする請求項2に記載のデータ処理装置。

【請求項4】 前記テキストを入力するタイプから前記候補を選択して入力するタイプに変換された入力欄が、元の前記テキストを入力するタイプとなるように、変換された前記データを逆変換する手段を備えることを特徴とする請求項2に記載のデータ処理装置。

【請求項5】 前記検出手段により検出された入力欄が、予め設定された複数の項目を有する項目リストの中から選択して入力する入力欄であった場合に、前記変換手段は、前記項目リストの項目または順位が前記認識候補に基づいて変更されるように、前記データを変換することを特徴とする請求項1に記載のデータ処理装置。

【請求項6】 前記音声認識手段は、複数の前記認識候補の各々の尤度を計算し、前記変換手段は、前記項目のリストが各候補の尤度の順番に表示されるように、前記データを変換することを特徴とする請求項5に記載のデータ処理装置。

【請求項7】 前記表示手段が前記項目のリストを表示する場合に、前記認識候補と一致する項目を他の項目よりも優先的な順番で表示するように、前記変換手段は前記データを変換することを特徴とする請求項5に記載のデータ処理装置。

【請求項8】 前記音声認識手段が、予め設定されている前記項目の中から前記認識候補を作成することを特徴

とする請求項5に記載のデータ処理装置。

【請求項9】 前記表示手段が前記項目のリストを表示する場合に、前記認識候補と一致する項目とそれ以外の項目とを識別可能に表示するように、前記変換手段は前記データを変換することを特徴とする請求項5に記載のデータ処理装置。

【請求項10】 前記変換手段は、項目リストから選択するタイプの入力欄に音声による入力が複数回あった場合は、当該リスト中で前回の認識候補と一致する項目の表示順位を下げるように、前記データを変換することを特徴とする請求項5に記載のデータ処理装置。

【請求項11】 前記変換手段は、項目リストから選択するタイプの入力欄に音声による入力が複数回あり、かつ、前回の認識候補と今回の認識候補に重複する候補があった場合は、当該重複する候補を今回の認識候補より削除するように、前記データを変換することを特徴とする請求項5に記載のデータ処理装置。

【請求項12】 前記音声認識処理において前記認識候補を作成できなかった場合に、音声認識に失敗した旨を報知する手段を備えたことを特徴とする請求項1乃至11のいずれかに記載のデータ処理装置。

【請求項13】 前記報知は、音声による報知であることを特徴とする請求項12に記載のデータ処理装置。

【請求項14】 前記変換手段が、前記音声認識手段が認識候補を作成できなかった場合に、前記データを、音声認識に失敗した旨が表示されるように変換することを特徴とする請求項1乃至11のいずれかに記載のデータ処理装置。

【請求項15】 所定のマークアップ言語で記述されたグラフィカルインターフェース用のデータの内容を表示する工程と、

前記データの表示内容に含まれる入力欄の1つに対して、音声により情報を入力をする工程と、

入力された前記音声に対して音声認識処理を実行し、その認識候補を作成する音声認識工程と、

前記表示内容に含まれる入力欄のうち、前記音声入力がなされた入力欄を検出する工程と、

検出した入力欄に対する前記データを、作成した前記認識候補に基づいて変換する工程と、を含むことを特徴とするデータ処理方法。

【請求項16】 所定のマークアップ言語で記述されたデータを処理するためのコンピュータプログラムを記録した記録媒体であって、該コンピュータプログラムが、前記マークアップ言語で記述されたグラフィカルインターフェース用のデータの内容を表示する工程のコードと、

前記データの表示内容に含まれる入力欄の1つに対して、音声により情報を入力をする工程のコードと、

入力された前記音声に対して音声認識処理を実行し、その認識候補を作成する音声認識工程のコードと、

前記表示内容に含まれる入力欄のうち、前記音声入力となされた入力欄を検出する工程のコードと、検出した入力欄に対する前記データを、作成した前記認識候補に基づいて変換する工程のコードと、を含むことを特徴とする記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、ネットワーク上で提供されるHTMLデータなどのマークアップ言語で記述されたデータの処理に関し、特に、ブラウザからの音声入力に関する。

【0002】

【従来の技術】情報検索、アンケート、商品の注文などを目的として利用できるWebページには、テキスト入力フォームや選択メニュー、ボタンなどを配置して、キーボードやマウスなどで操作可能なGUI(Graphical User Interface)を用意したものが多く見られる。

【0003】一方、近年、キーボードやマウスなどに加えて、音声が入力インターフェースとして用いられるようになってきた。例えば、ブラウザ上に表示されたWebページのテキスト入力フォームへ音声で情報を入力したり、選択メニューから項目を音声で選択する方法がある。これらの方法では、テキスト入力フォームには、最も尤度の高い認識結果を入力し、選択メニューでは、最も尤度の高い項目を選択している。しかし、音声認識で100%の認識率を達成するのは困難である。そこで、複数の認識候補を画面上で、音声入力したテキスト入力フォーム、あるいは、選択メニューとは別の領域に表示し、その中から利用者に正しい結果を選択させ、テキスト入力フォームでは選択された結果を入力し、選択メニューでは選択された結果の項目を選択する方法等が採用されている。

【0004】

【発明が解決しようとする課題】しかし、従来の手法では、音声入力における複数の認識候補が、ブラウザ上で表示されたWebページのテキスト入力フォームや選択メニューとは別の領域に表示されるため、使い勝手が必ずしもよくなかった。

【0005】従って、本発明の目的は、ブラウザ上で表示されたWebページのテキスト入力フォームや選択メニュー上で、音声入力における複数の認識候補を表示可能なデータ処理装置、データ処理方法、及び、記録媒体を提供することにある。

【0006】

【課題を解決するための手段】本発明によれば、所定のマークアップ言語で記述されたグラフィカルインターフェース用のデータの内容を表示する表示手段と、前記表示手段による表示内容に含まれる入力欄の1つに対して、音声により情報を入力するための音声入力手段と、入力された前記音声に対して音声認識処理を実行

し、その認識候補を作成する音声認識手段と、前記表示内容に含まれる入力欄のうち、前記音声入力手段により入力となされた入力欄を検出する検出手段と、前記検出手段により検出した入力欄に対する前記データを、作成した前記認識候補に基づいて変換する手段と、を備えたことを特徴とするデータ処理装置が提供される。

【0007】また、本発明によれば、前記検出手段により検出された入力欄が、該入力欄にテキストを入力するタイプの入力欄であった場合に、前記変換手段は、前記入力欄が、作成した前記認識候補を選択候補とする選択リストから候補を選択して入力するタイプとなるように、前記データを変換することを特徴とするデータ処理装置が提供される。

【0008】また、本発明によれば、前記検出手段により検出された入力欄が、予め設定された複数の項目を有する項目リストの中から選択して入力する入力欄であった場合に、前記変換手段は、前記項目リストの項目または順位が前記認識候補に基づいて変更されるように、前記データを変換することを特徴とするデータ処理装置が提供される。

【0009】また、本発明によれば、前記音声認識手段が、予め設定されていた前記項目の中から前記認識候補を作成することを特徴とするデータ処理装置が提供される。

【0010】また、本発明によれば、所定のマークアップ言語で記述されたグラフィカルインターフェース用のデータの内容を表示する工程と、前記データの表示内容に含まれる入力欄の1つに対して、音声により情報を入力をする工程と、入力された前記音声に対して音声認識処理を実行し、その認識候補を作成する音声認識工程と、前記表示内容に含まれる入力欄のうち、前記音声入力となされた入力欄を検出する工程と、検出した入力欄に対する前記データを、作成した前記認識候補に基づいて変換する工程と、を含むことを特徴とするデータ処理方法が提供される。

【0011】また、本発明によれば、所定のマークアップ言語で記述されたデータを処理するためのコンピュータプログラムを記録した記録媒体であって、該コンピュータプログラムが、前記マークアップ言語で記述されたグラフィカルインターフェース用のデータの内容を表示する工程のコードと、前記データの表示内容に含まれる入力欄の1つに対して、音声により情報を入力をする工程のコードと、入力された前記音声に対して音声認識処理を実行し、その認識候補を作成する音声認識工程のコードと、前記表示内容に含まれる入力欄のうち、前記音声入力となされた入力欄を検出する工程のコードと、検出した入力欄に対する前記データを、作成した前記認識候補に基づいて変換する工程のコードと、を含むことを特徴とする記録媒体が提供される。

【0012】

【発明の実施の形態】以下、本発明の好適な実施の形態について説明する。

＜実施形態1＞図1は、本発明が機能的に実現され得るハードウェアの構成例を示した図である。

【0013】CPU101は、全体を統括制御するものであり、ROM102に格納されているプログラムをRAM103に読み出し、その読み出したプログラムに基づいて、各種処理動作を実行する。

【0014】ROM102は、CPU101で実行する処理の各種プログラムを格納している。RAM103は、ROM102に格納されている各種プログラムの実行に必要な記憶領域を提供する。

【0015】二次記憶装置104は、OSや各種プログラムを格納している。ただし、パーソナルコンピュータなどの汎用装置ではなく、専用装置で構成する場合には、ROM102内にOSや各種プログラムを格納しても構わない。この格納されたプログラムをRAM103に読み出すことによってCPU101が処理を実行することが可能である。また、二次記憶装置としては、ハードディスク装置、フロッピー（登録商標）ディスクドライブ、CD-ROMドライブなどを用いることができる。つまり、記憶媒体は如何なるものであっても構わない。

【0016】入力装置105としては、マウスやキーボード、又は、音声入力のためのマイク等を挙げることができる。これら複数の機器を同時に接続して使用できる。

【0017】出力装置106としては、HTMLデータの内容を表示するディスプレイ（CRTやLCDなど）や、処理の内容を報知するスピーカ等から構成される。

【0018】ネットワークI/F107は、ネットワーク上で提供されるHTMLデータを取得するためのものである。バス108は、上記の各構成を接続する。

【0019】図2は、本発明の一実施形態に係るブラウザ装置の機能ブロック図であり、図1に示したハードウェアを用いることにより機能させることができる。

【0020】207は音声入力処理部であり、HTMLデータの内容を表示するブラウザ上に含まれる入力欄、すなわち入力フォームに、ユーザが情報の入力又は項目の選択を行うために発声した音声を計算機上に取り込むものであり、図中、音声情報201は入力された音声情報を示している。また、音声入力があると、音声入力された入力フォーム（例えば、フォーカスのあっている入力フォーム）の識別子が取得され、保存される。図中、203は、取得し、保存された入力フォームの識別子を示している。音声入力処理部207は、入力装置105（マイクなど）からの音声入力の他、二次記憶装置104に記憶されている音声データファイル、ネットワークI/F107などからの音声情報を受け付けることもできる。音声データの形式は特に限定されるものではない。

【0021】208は音声認識部であり、音声情報201に対して音声認識処理を実行し、その認識候補を作成する。図中、202は、認識候補を示している。認識候補202は、音声認識の確からしさを示す尤度の高い順に複数得ることもできる。

【0022】209は、入力フォーム処理部であり、HTMLデータの入力フォーム情報を取得したり、格納したりする。入力フォーム情報（現在値）204は、入力フォームの識別子や、入力フォームの設定値などである。設定値とは、テキスト入力タイプの入力フォームでは、入力されたテキストであり、選択候補から1つを選択するタイプの入力フォームでは、選択された項目のことである。なお、“（現在値）”とあるのは、現在ブラウザ装置が表示中の入力フォームの設定値であることを意味する。

【0023】211は、HTMLデータ処理部であり、表示されているページが異なるページになった時に、新たなHTMLデータ206を取得し、保存する。また、入力フォーム情報（初期値）205を取得したり、保存したりする。入力フォーム情報（初期値）205は、入力フォームの識別子、入力フォームのタイプ（テキスト入力、選択候補からの1つの選択など）などである。“（初期値）”としているのは、後で説明するようにHTMLデータが変換される前のデータであることを示している。

また、HTMLデータ処理部211は、HTMLデータ206がHTMLデータ作成部210で変更された時に、HTMLデータ206をブラウザにPushして、ディスプレイに表示させる。

【0024】210はHTMLデータ作成部であり、認識候補202、音声入力された入力フォームの識別子203、入力フォーム情報（現在値）204、入力フォーム情報（初期値）205、HTMLデータ206等に基づいて、HTMLデータ206を変換、保存等する。

【0025】次に、HTMLデータ作成部210の動作をフローチャートを用いて説明する。なお、本実施形態では、音声入力される入力フォームは、テキスト入力タイプと選択候補からの1つの選択のタイプのみであるとして説明する。

【0026】ステップS300では、HTMLデータ206から入力フォームの記述を検出し、見つけた入力フォームを現在処理対象としている入力フォームとする。

【0027】ステップS301では、音声入力された入力フォームの識別子203と現在処理対象としている入力フォームの識別子とを比較し、現在処理対象としている入力フォームに対して、音声入力されたか否か判定する。現在処理対象としている入力フォームに音声入力されていないと判定した時は、ステップS305に進む。

【0028】ステップS305では、入力フォーム情報（現在値）204から、現在処理対象としている入力フ

10

20

30

40

50

フォームに対応する入力フォームの設定値を取得し、現在処理対象としている入力フォームの設定値と比較する。入力フォームの設定値が異なっていた場合には、現在処理対象としている入力フォームの設定値を、前記取得した設定値に変更する。

【0029】一方、ステップS301で、現在処理対象としている入力フォームが音声入力された入力フォームであると判定された時は、ステップS302に進む。

【0030】ステップS302では、入力フォーム情報（初期値）205より、現在処理対象としている入力フォームのタイプを取得する。そして、取得したタイプに応じて、テキスト入力タイプの場合にはステップS303に進み、選択候補からの1つの選択タイプの場合にはステップS304に進む。

【0031】ステップS303では、現在対象としている入力フォーム（テキスト入力タイプの入力フォーム）を、選択候補からの1つの選択タイプの入力フォームに変更し、認識候補202を埋め込む。認識候補として、音声認識の確からしさを示す尤度の高い順に複数得られる場合には、尤度が高い候補が表示される優先順位を高くするように記述を変更する。例えば、尤度が高いものほど選択候補の上位に並べる方法がある。

【0032】例えば、現在処理対象としている入力フォームの記述が、

```
<INPUT type="text" name="word">
```

であった時（ブラウザ上での表示例を図4（a）に示す。）に、認識候補として、尤度順に“中山”、“和歌山”、“高山”、“岡山”と得られた場合には、

```
<SELECT name="word">
```

```
<OPTION SELECTED> 中山
```

```
<OPTION> 和歌山
```

```
<OPTION> 高山
```

```
<OPTION> 岡山
```

```
</SELECT>
```

と記述を変換する（ブラウザ上での表示例を図4（b）に示す。）。なお、CGI等で、サーバへ返送する変数であるname="xxx"の部分はそのままとする。

【0033】ここでは、第一位の認識候補に対しては、OPTIONタグの中でSELECTEDを記述したが、SELECTEDの記述が1つも存在しない時に、ブラウザがデフォルトで選択候補の最上位にあるものを設定値としてもつ時には、SELECTEDの記述は省略しても良い。

【0034】なお、認識候補が1つしかない場合には、選択候補からの1つの選択タイプのフォームに変更せずに、テキスト入力タイプのフォームにしても良い。

【0035】例えば、認識候補が“中山”だけの場合には、

```
<SELECT name="word">
```

```
<OPTION SELECTED> 中山
```

```
</SELECT>
```

と記述を変換しても良いし、

```
<INPUT type="text" name="word" value="中山">
```

と記述を変換しても良い（ブラウザ上での表示例を図4（c）に示す。）。

【0036】また、認識候補が1つもなかった場合には、認識に失敗したことをユーザに知らせるために、“認識失敗”などを埋め込むように記述を変換しても良いし（ブラウザ上での表示例を図4（d）に示す。）、何も埋め込まないで、なんらかの方法によってユーザに認識が失敗したことを提示しても良い。例えば、スピーカから、“認識に失敗しました”と音声を出力すれば良い。

【0037】ただし、SELECTタグの中には最低1つのOPTIONタグが必要なため、何も埋め込まない場合は、選択候補から1つの選択タイプは使用できず、テキスト入力タイプのフォームにする必要がある。

【0038】一方、ステップS304では、現在対象としている入力フォーム（予め設定された項目である選択候補から1つを選択するタイプの入力フォーム）において、予め設定された項目である選択候補中で、認識候補202と一致する選択候補が表示される優先順位を高くするように記述を変換する。例えば、予め設定された項目である選択候補中、認識候補と一致する選択候補を最上位に並べ、それ以外の選択候補を下位に並べる方法がある。なお、ここでは、予め設定された項目である選択候補以外の語が認識候補として得られることはないものとして説明した。認識候補として、音声認識の確からしさを示す尤度の高い順に複数得られる場合には、尤度が高い候補が表示される優先順位を高くするように記述を変更する。例えば、尤度が高いものほど選択候補の上位に並べる方法がある。

【0039】例えば、現在処理対象としている入力フォームの記述が、

```
<SELECT name="city">
```

```
<OPTION> 東京
```

```
<OPTION> 横浜
```

```
<OPTION> 大阪
```

```
<OPTION> 岡山
```

```
<OPTION> 名古屋
```

```
:
```

```
:
```

```
:
```

```
</SELECT>
```

であった時（ブラウザ上での表示例を図5に示す。）に、認識候補として、尤度順に“中山”、“和歌山”、“高山”、“岡山”と得られた場合には、

```
<SELECT name="city">
```

```
<OPTION SELECTED> 中山
```

```
<OPTION> 和歌山
```

```
<OPTION> 高山
```

```
<OPTION> 岡山
```

<OPTION> 東京
 <OPTION> 横浜
 <OPTION> 大阪
 <OPTION> 名古屋
 :
 :
 :

</SELECT>

と記述を変更する（ブラウザ上での表示例を図6に示す）。

【0040】ここでは、第一位の認識候補に対しては、OPTIONタグの中でSELECTEDを記述したが、SELECTEDの記述が1つも存在しない時に、ブラウザがデフォルトで選択候補の最上位にあるものを設定値としてもつ時には、SELECTEDの記述は省略しても良い。

【0041】なお、認識候補が1つしかない場合には、並べ替えを行わず、認識結果のOPTIONタグをSELECTEDにするのみでも良い。

【0042】例えば、認識候補が“岡山”だけの場合には、

```
<SELECT name="city">
<OPTION> 東京
<OPTION> 横浜
<OPTION> 大阪
<OPTION SELECTED> 岡山
<OPTION> 名古屋
:
:
:
```

</SELECT>

と記述を変換しても良い（ブラウザ上での表示例を図7に示す。）。

【0043】認識候補が1つもなかった場合には、認識に失敗したことをユーザに知らせるために、“認識失敗”などを埋め込むように記述を変換してよい。認識失敗を埋め込む位置は、最上位でもよいし（ブラウザ上での表示例を図8に示す。）、最下位でもよい（ブラウザ上での表示例を図9に示す。）。

【0044】また、並べ替えを行わず、何も埋め込まないで、なんらかの方法によってユーザに認識が失敗したことを提示しても良い。例えば、スピーカから、“認識に失敗しました”と音声を出力すれば良い。

【0045】ステップS306では、全入力フォームに対して処理を終了したか否か判定し、終了していた場合には、HTMLデータの変換作業を終了する。終了していない場合には、ステップS300に戻り、処理を繰り返す。

<実施形態2>本実施形態は、実施形態1に対して、以下の処理を追加したものである。

【0046】テキスト入力タイプの入力フォームに対し

ては、周りの雑音などが大きく、認識率が低い場合には、音声入力ではなく、キーボードなどの入力装置を使用の方がユーザにとって望ましい場合がある。

【0047】そのため、テキスト入力タイプの入力フォームが、選択候補からの1つの選択タイプの入力フォームに変換されていた場合に、もとのテキスト入力タイプの入力フォームに変換する（戻す）手段を用意する。

【0048】具体的には、入力フォーム情報（現在値）に、入力フォームのタイプを持ち、ユーザからの要求があった時に、入力フォーム情報（初期値）の入力フォームのタイプと比較して、入力フォーム情報（初期値）の入力フォームのタイプがテキスト入力タイプで、かつ、入力フォーム情報（現在値）の入力フォームのタイプが選択候補からの1つの選択の場合に、テキストタイプの入力フォームに記述を変換する。また、入力フォーム情報（現在値）に、テキスト入力フォームから選択候補からの1つの選択タイプの入力フォームに変換されたか否かのフラグを追加し、ユーザからの要求があった時に、このフラグによって判定し、テキスト入力フォームから選択候補からの1つの選択タイプの入力フォームに変換された場合に、テキストタイプの入力フォームに記述を変換しても良い。ユーザからの要求があった時ではなく、同じテキストタイプの入力フォームに、一定回数以上音声入力がなされた時に変換しても良い。

<実施形態3>本実施形態は、実施形態1に対して、以下の処理を追加したものである。

【0049】選択タイプの入力フォームの記述の変換において、同じ入力フォームに対して、複数回音声入力があった場合も、上述した処理をそれぞれ実行すれば足りるが、この場合、音声入力をし直すということは、選択候補の上位に配置した認識候補の中に正解がなかったと推定できる。ゆえに、前回の認識候補と一致する選択候補が表示される優先順位を下げるようにHTMLの記述を変換する。例えば、前回の認識候補と一致する選択候補をリストの最下位に配置するようにする。

【0050】具体的には、例えば、バッファを2つ持ち、前回の認識候補と今回の認識候補を保持する。そして、前回の認識候補を最下位に配置し、今回の認識候補を最上位に配置する。これら以外の選択項目をその間に配置する。

【0051】前回の認識候補と今回の認識候補に重複するものがあった場合には、前回の認識候補を最下位に配置し、今回の認識候補から重複するものを除いたものを最上位に配置すれば良い。

【0052】そして、HTMLデータの変換処理が終わった後で、今回の認識候補を前回の認識候補とする。また、異なるページが表示された時に、クリアする。

【0053】例えば、現在処理対象としている入力フォームの記述が、

```
<SELECT name="city">
```

```

<OPTION> 東京
<OPTION> 横浜
<OPTION> 大阪
<OPTION> 岡山
<OPTION> 名古屋
:
:
:

```

```

</SELECT>

```

であり（ブラウザ上での表示例を図5に示す。）、認識候補として、尤度順に“中山”、“和歌山”、“高山”、“岡山”と得られた場合に、

```

<SELECT name="city">
<OPTION SELECTED> 中山
<OPTION> 和歌山
<OPTION> 高山
<OPTION> 岡山
<OPTION> 東京
<OPTION> 横浜
<OPTION> 大阪
<OPTION> 名古屋
:
:
:
</SELECT>

```

と記述が変換されているとする（ブラウザ上での表示例を図6に示す。）。この時、音声入力し直した結果、認識候補として、尤度順に、“徳山”、“館山”、“高山”と得られた場合に、

```

<SELECT name="city">
<OPTION SELECTED> 徳山
<OPTION> 館山
<OPTION> 東京
<OPTION> 横浜
<OPTION> 大阪
<OPTION> 名古屋
:
:
:

```

```

<OPTION> 中山
<OPTION> 和歌山
<OPTION> 高山
<OPTION> 岡山
</SELECT>

```

と記述を変換すればよい（ブラウザ上での表示例を図10、図11に示す。）

＜実施形態4＞本実施形態は、実施形態1、あるいは、実施形態3に以下の処理を追加したものである。

【0054】選択タイプの入力フォームの記述の変換において、認識候補と一致する選択候補と、それ以外の選

択候補とが区別して表示されるように、記述を変換する。例えば、選択候補中、認識候補と一致する選択候補を上位に並べ、それ以外の選択候補を下位に並べている場合には、両者の間に、空の項目を追加し、区別を容易にすることも可能である。

【0055】例えば、現在対象としている入力フォームの記述が、

```

<SELECT name="city">
<OPTION> 東京
<OPTION> 横浜
<OPTION> 大阪
<OPTION> 岡山
<OPTION> 名古屋
:
:
:
</SELECT>

```

であり（ブラウザ上での表示例は図5）、認識候補として、尤度順に“中山”、“和歌山”、“高山”、“岡山”と得られた場合に、

```

<SELECT name="city">
<OPTION SELECTED> 中山
<OPTION> 和歌山
<OPTION> 高山
<OPTION> 岡山
<OPTION>
<OPTION> 東京
<OPTION> 横浜
<OPTION> 大阪
<OPTION> 名古屋
:
:
:
</SELECT>

```

と記述を変換すればよい（ブラウザ上での表示例を図12に示す。）。

＜実施形態5＞本実施形態は、実施形態1、実施形態3、あるいは、実施形態4の実施形態に以下の処理を追加したものである。

40 【0056】選択タイプの入力フォームへの音声入力の場合に、認識候補として、選択一覧以外の語が得られる場合には、選択候補以外の語が得られる場合には、認識候補と選択候補を比較し、選択候補に含まれない認識候補を削除し、選択候補に含まれる認識候補のみを使用して、HTMLデータの変換を行うこともできる。

【0057】以上説明した実施形態では、変換対象となるデータをHTMLデータとして説明したが、それに限らず、グラフィカルユーザインターフェース用の表示データとして他のマークアップ言語で作成されたデータにも適用可能である。

【0058】以上説明した実施形態によれば、ユーザが入力しようとしているブラウザ上の入力フォームに、音声入力における認識候補を表示し、GUI等で選択することが出来るため、より自然なインタフェースになる。

【0059】また、選択候補から1つを選択するタイプの入力フォームにおいて、選択候補が多くある場合には、GUIなどでの選択に時間がかかるが、以上説明した実施形態によれば、音声入力における認識候補が尤度順に選択候補の上位に配置されるので、認識候補の中に正解がある場合には、第一位でなくても上位にあるため、GUIにより短時間で選択でき、かつ、第二位以下の認識候補に正解があった場合に、発声し直す必要がなくなる。また、同音異表記の認識候補があった場合には、GUIを用いて選択できるようになる。

【0060】以上、本発明の好適な実施形態について説明したが、本発明の目的は、前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記憶媒体（または記録媒体）を、システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ（またはCPUやMPU）が記憶媒体に格納されたプログラムコードを読み出し実行することによっても、達成されることは言うまでもない。この場合、記憶媒体から読み出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記憶した記憶媒体は本発明を構成することになる。また、コンピュータが読み出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているオペレーティングシステム(OS)などが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0061】さらに、記憶媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張カードやコンピュータに接続された機能拡張ユニットに備わるメモリに書込まれた後、そのプログラムコードの指示

に基づき、その機能拡張カードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0062】

【発明の効果】以上述べてきた通り、本発明によれば、ブラウザ上の入力フォームに音声入力における認識候補を表示し、選択することができる。

【図面の簡単な説明】

10 【図1】本発明が機能的に実現され得るハードウェアの構成例を示した図である。

【図2】本発明の一実施形態に係るブラウザ装置の機能ブロック図である。

【図3】第一の実施形態のHTMLデータ作成部の動作を表すフローチャートである。

【図4】(a)乃至(d)は、テキスト入力タイプの入力フォームの記述を変換した際のブラウザの表示例を示した図である。

20 【図5】選択タイプの入力フォームの記述を変換した際のブラウザの表示例を示した図である。

【図6】選択タイプの入力フォームの記述を変換した際のブラウザの表示例を示した図である。

【図7】選択タイプの入力フォームの記述を変換した際のブラウザの表示例を示した図である。

【図8】選択タイプの入力フォームの記述を変換した際のブラウザの表示例を示した図である。

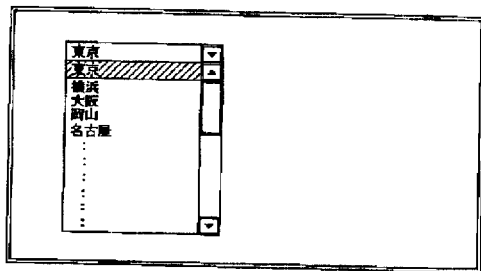
【図9】選択タイプの入力フォームの記述を変換した際のブラウザの表示例を示した図である。

30 【図10】複数回の音声入力があった場合における選択タイプの入力フォームの記述を変換した際のブラウザの表示例を示した図である。

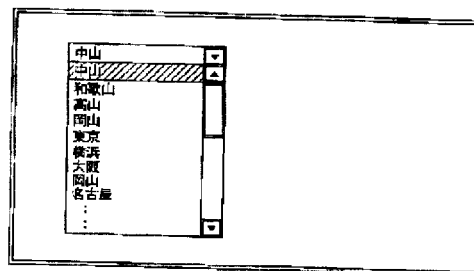
【図11】複数回の音声入力があった場合における選択タイプの入力フォームの記述を変換した際のブラウザの表示例を示した図である。

【図12】選択するタイプの入力フォームの記述を変換した際のブラウザの表示例を示した図である。

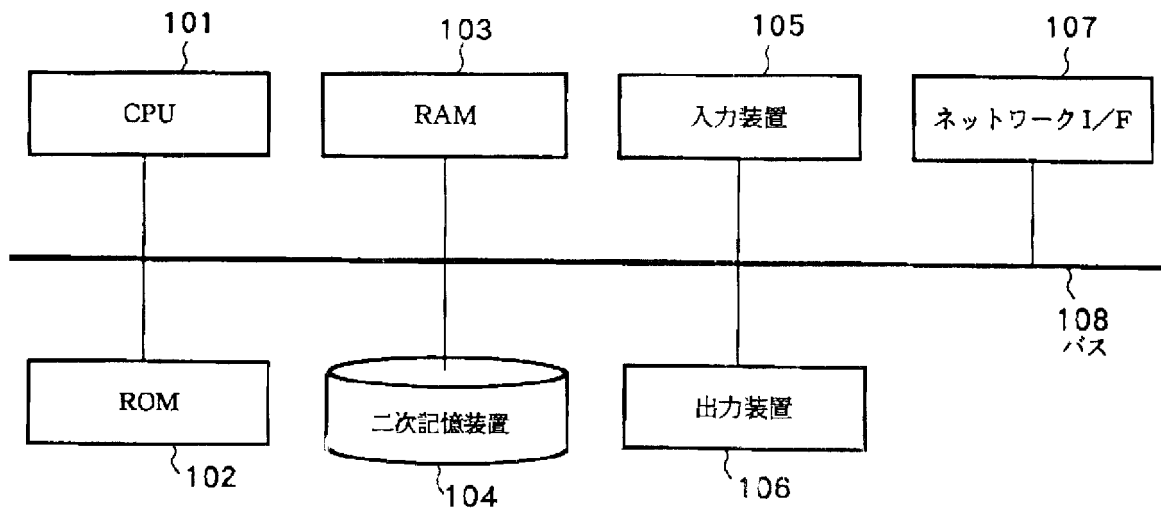
【図5】



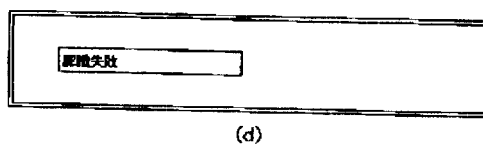
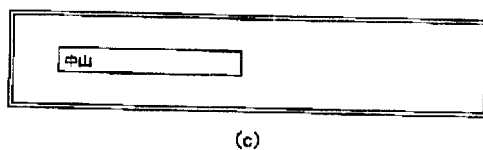
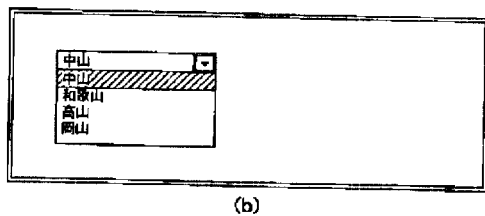
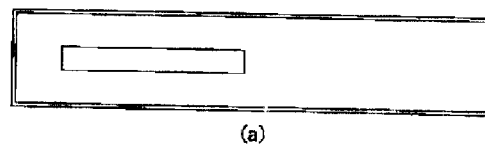
【図6】



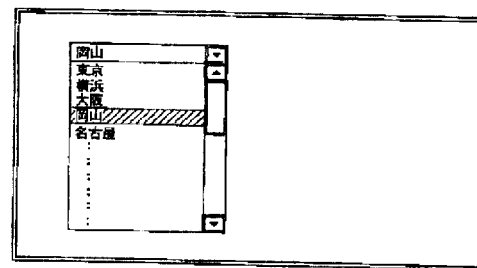
【図1】



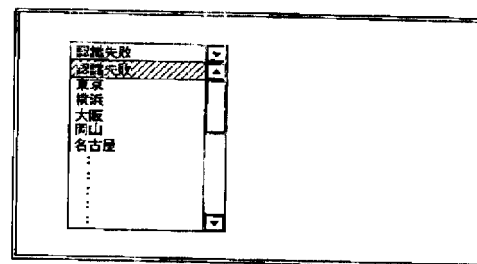
【図4】



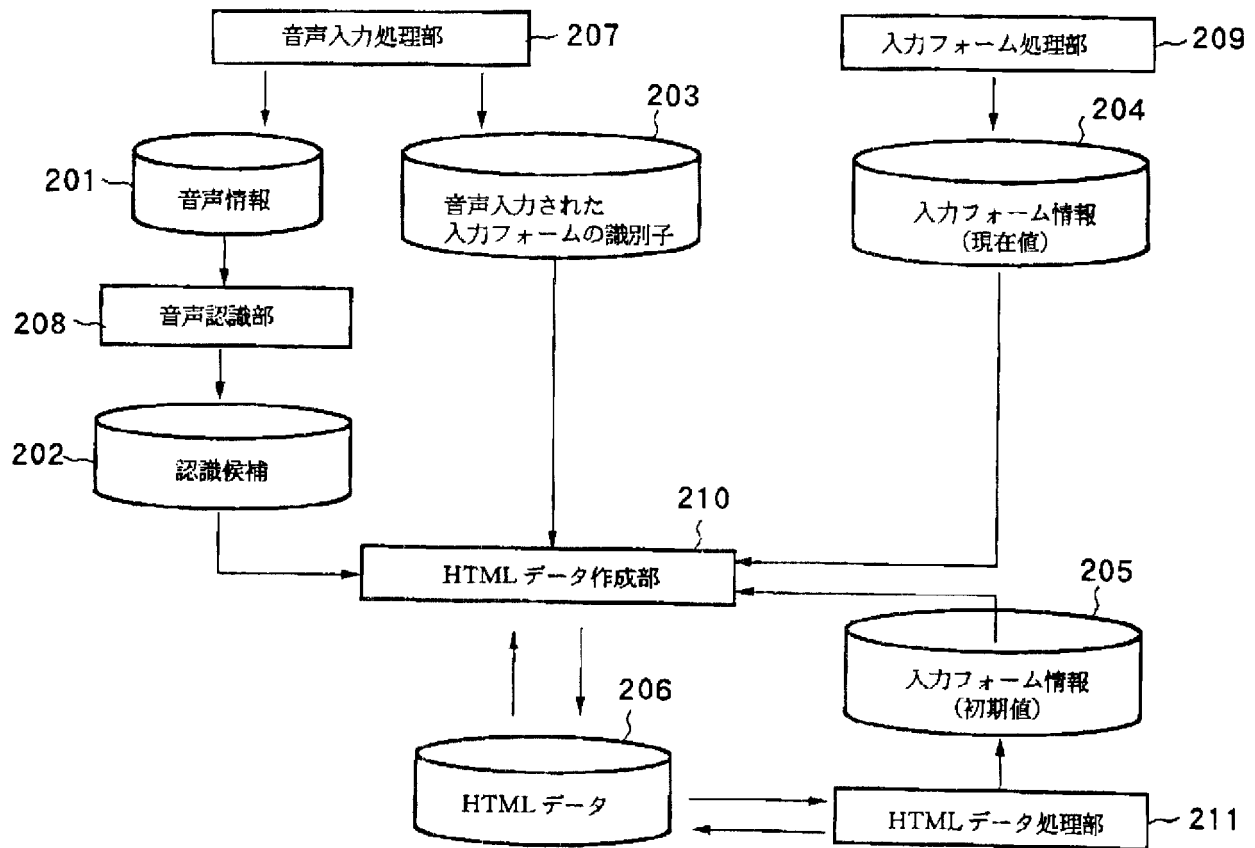
【図7】



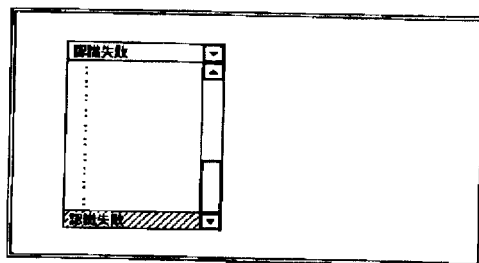
【図8】



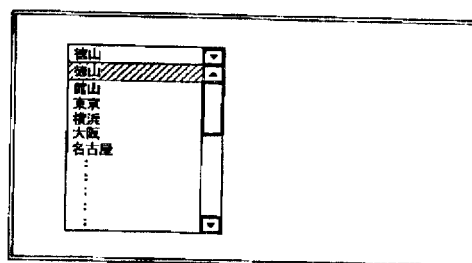
【図2】



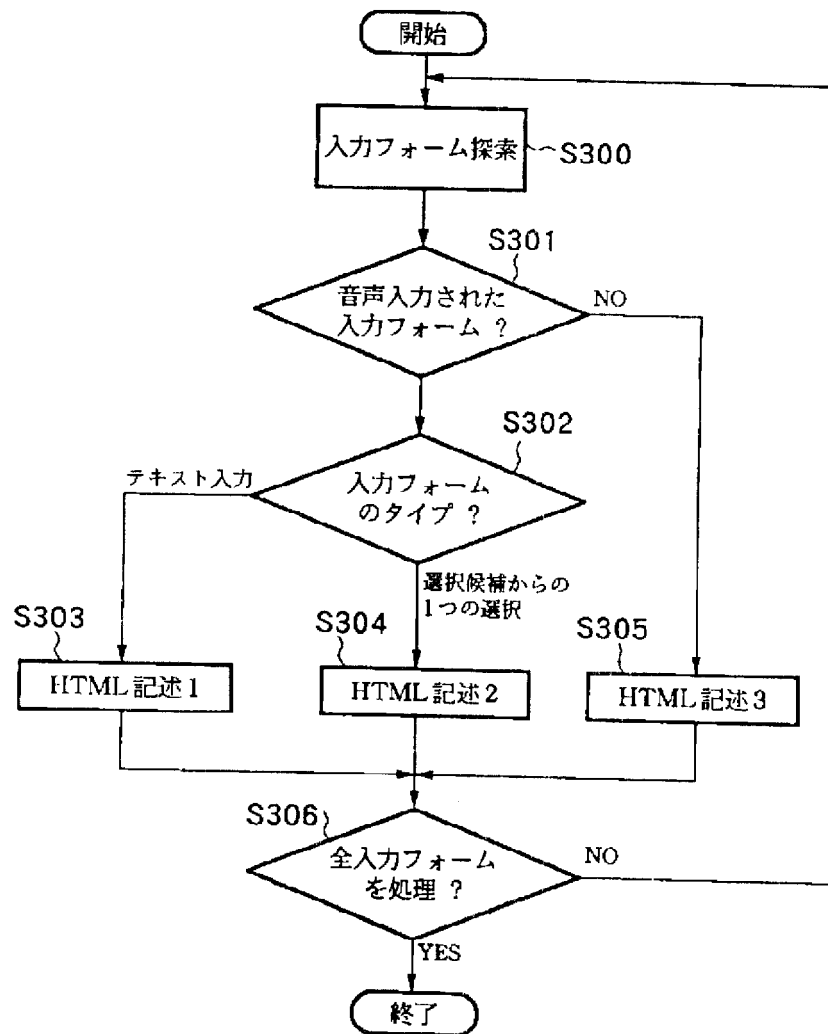
【図9】



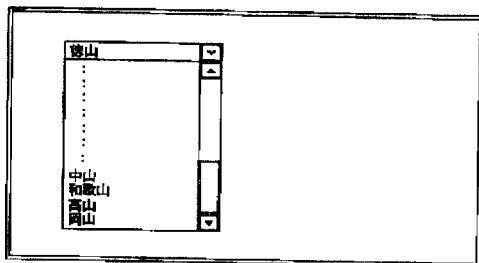
【図10】



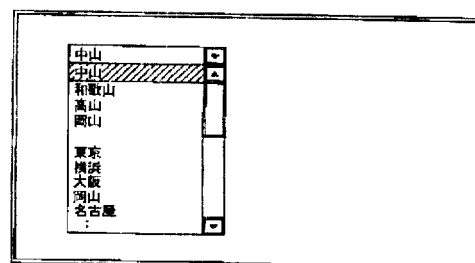
【図3】



【図11】



【図12】



フロントページの続き

(51) Int. Cl.⁷

識別記号

F I
G 1 0 L 3/00

テマコード (参考)

5 6 1 D

5 7 1 R

(72)発明者 伊藤 史朗

東京都大田区下丸子3丁目30番2号 キヤ
ノン株式会社内

Fターム(参考) 5D015 KK02 LL05 LL06 LL08